

САНКТ-ПЕТЕРБУРГСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
КАФЕДРА КОМПЬЮТЕРНОГО МОДЕЛИРОВАНИЯ  
И МНОГПРОЦЕССОРНЫХ СИСТЕМ

**Забелина Татьяна Сергеевна**

**Выпускная квалификационная работа бакалавра**

**Методы и инструменты  
3D-реконструкции изображений лиц**

Направление 010300

Фундаментальная информатика и информационные технологии

Научный руководитель,  
доктор физ.-мат. наук,  
профессор  
Андрианов С. Н.

Санкт-Петербург  
2017

# Оглавление

<b>Введение</b>	<b>3</b>
<b>1. Обзор методов трехмерной реконструкции</b>	<b>9</b>
1.1. Основные принципы 3D-реконструкции . . . . .	9
1.2. 3D Face Morphable Model . . . . .	12
1.3. Меры качества реконструкции для подгонки к изображениям	15
<b>2. Реализация</b>	<b>19</b>
2.1. Используемые технологии . . . . .	19
2.2. Библиотека eos . . . . .	19
2.3. Описание алгоритма . . . . .	20
<b>3. Эксперименты</b>	<b>23</b>
<b>Заключение</b>	<b>26</b>

# Введение

## Трехмерная реконструкция и ее применение

В последние десятилетия вычислительная техника проникла во множество сфер жизнедеятельности человека. Технологические инновации позволяют проводить научные исследования, облегчают и автоматизируют ручной труд, открывают перед нами совершенно новые и привлекательные возможности. Компьютерное зрение – относительно молодое, но активно развивающееся направление с огромной сферой применения. Так, решение огромного количества задач современного мира производится посредством обработки цифровых изображений. Однако для многих задач уже не хватает плоских изображений, и множество предприятий нуждаются в трехмерных моделях тех сцен, с которыми они работают.

Цифровые 3D модели часто создают вручную, когда художник или инженер в специальном редакторе создает трехмерную модель. Однако этот способ требует достаточно много времени и умений. А в некоторых ситуациях требуется получение трехмерных моделей в режиме реального времени, и тогда даже самые умелые человеческие руки не в состоянии справиться с задачей. Для роботизации и автоматизации построения 3D моделей все чаще применяются методы **трехмерной реконструкции** – воссоздания трехмерных поверхностей наблюдаемых объектов. Реконструирование 3D моделей в условиях производства значительно снижает временные и материальные затраты на создание моделей, способствует повышению их точности в сравнении с ручным моделированием.

Потребности построения 3D моделей возникают во многих областях [19]:

- *медицина* – построение моделей органов и частей тела производится для обнаружения отклонений и определения необходимости лечения;
- *культура и история* – производится реконструкция объектов архитектуры и искусства, а также находок на местах раскопок, что позволяет предоставлять подробную информацию для исследователей [7, 4];
- *промышленное производство* – компьютерные макеты деталей механизмов создаются практически в каждом предприятии, а современные методы реконструкции позволяют построить мощные системы контроля качества при поточном производстве и проверять производимые детали на соответствие проектным параметрам в реальном времени;
- *индустрия развлечений* – качественная реконструкция поверхностей реальных объектов способствует созданию более естественных персонажей и окружений в 3D-кинофильмах, компьютерных играх, и при моделировании виртуальной реальности;
- *системы компьютерного зрения* – основной задачей таких систем является обнаружение, отслеживание и классификация объектов окружающего пространства; трехмерная реконструкция применяется для решения задач построения карт местности и прокладки маршрута, идентификации объектов, отслеживания их взаимного расположения и перемещения, выделения объектов и их признаков для извлечения дополнительных знаний.

### 3D-реконструкция лиц

Методы трехмерной реконструкции широко используются для восстановления 3D-геометрии человеческих лиц. Восстановление 3D поверхностей и трехмерное распознавание лиц применимо при решении следующих задач [16]:

- *обеспечение безопасности* – с помощью камер видеонаблюдения в местах массового пребывания людей можно отслеживать нахождение от-



дельных личностей, в том числе преступников;

- *верификация и идентификация* пользователей различных систем;
- *проведение контроля* на пограничных паспортно-визовых контрольных пунктах, проходных и КПП режимных, стратегических и коммерческих объектов;
- *бесконтактный контроль психоэмоционального состояния человека* для предотвращения несчастных случаев (в случае наблюдения за диспетчерами, охранниками, пилотами, машинистами, водителями), а также для выявления агрессивных и потенциально опасных людей, ориентированных на совершение преступлений и террористических актов, с целью обеспечения безопасности в аэропортах, вокзалах и других особо охраняемых объектах;
- *реклама и маркетинговые исследования* – интерактивные сервисы, реагирующие на мимику покупателя, позволяют получать оценку удовлетворенности потребителя, а также таргетированная реклама позволяет предсказывать, какой продукт может быть действительно интересен конкретному пользователю.

Трехмерное распознавание лиц осуществляется по реконструированным поверхностям, поэтому получение 3D-геометрии лица является актуальным направлением исследований.

Согласно [21], алгоритмы трехмерного распознавания лиц используют информацию о глубине и кривизне поверхности, в то время как системы 2D-распознавания, традиционно используют признаки, основанные на цветовых характеристиках пикселей изображения. Системы 2D-распознавания весьма чувствительны к освещению и при неравномерном освещении лица достоверность распознавания заметно падает. Поэтому основным преимуществом трехмерного распознавания лиц перед двумерным является отсутствие зависимости от условий освещения [23]. Для систем 3D-распознавания изменения в освещении влияют лишь на текстуру

лица, а реконструированная поверхность лица сохраняется неизменной. Однако нельзя сказать, что 3D-распознавание лиц совсем не зависит от освещения. В зависимости от алгоритма реконструкции некоторые части лица (например, засвеченные участки либо области с очень низким контрастом) могут выглядеть как пробелы или выбросы (артефакты) на поверхности реконструкции. Устранение подобных артефактов в малых количествах производится вручную, а при увеличении их числа разрабатываются и используются специальные методы, например метод обратного расстояния [24].

Другим важным отличием технологий трехмерного распознавания по сравнению с двумерным является устойчивость к изменению позы [23]. Допустимое отклонение ракурса от фронтального положения составляет 15 градусов по вертикали и горизонтали для лучших образцов систем 2D-распознавания. При 3D-распознавании допустимый угол отклонения головы от фронтального ракурса может достигать 45 градусов. Более того, реконструированную трехмерную модель всегда можно привести программными средствами к виду фас для последующего распознавания стандартными двумерными алгоритмами.

В системах распознавания лиц используются устойчивые антропометрические точки (*landmarks*), расположение которых характеризует индивидуальные особенности лица. На 3D-моделях антропометрические точки определяются с большей точностью, чем на 2D-изображениях. Кроме того, точки на 3D-моделях имеют три координаты и, соответственно, дают больше информации, чем те же точки на двумерном изображении.

Из достоинств систем трехмерного распознавания лиц кроме их эффективности можно выделить также удобство в использовании [22]. Во-первых трехмерное распознавание является бесконтактным, что удобно как пользователям системы, так и ее владельцам. Во-вторых, данный способ распознавания не требует от человека специальных действий, так как часто направлен на распознавание лиц в режиме реального времени посредством обработки видеопотока камеры. Человеку не нужно позиционировать себя перед считывающим устройством, направлять взгляд, либо

сканировать палец или ладонь. Зачастую человек вообще может не обращать внимания на систему распознавания, а иногда даже и не знать о ее существовании. В-третьих, для распознавания не требуется дополнительное сканирование биометрических признаков человека. Базы данных с фотографиями пользователей присутствуют на многих объектах еще до установки биометрической системы.

Одним из важнейших недостатков трехмерного распознавания является высокая стоимость используемого оборудования, поскольку система 3D-распознавания требует гораздо больших вычислительных ресурсов, чем двумерные аналоги.

Получение трехмерной модели лица обычно предполагает наличия специального устройства, как, например, лазерный сканер, который измеряет время отражения лазера от объекта и получает информацию о глубине расположения точек на его поверхности. Такие приборы сложно использовать вне специальных лабораторий из-за трудностей в оборудовании необходимого окружения и их высокой цены. Лазерное сканирование показывает очень высокие результаты, однако даже при идеальном освещении сенсорные системы подвержены недостаткам: частым случаем является появление на модели артефактов в виде «дыр» и «выступов» из-за отсутствующих данных и ошибок восстановления.

В более простом случае реконструирование 3D моделей лиц производится по одной или нескольким фотографиям (фотограмметрия). При таком подходе, как правило, используют базы шаблонных моделей для поиска максимально близких к исходному изображению ключевых областей.

## Постановка задачи

На основе проведенного анализа целью данной работы является подробное изучение существующих методов 3D-реконструкции лиц, а также использование свободных инструментов для реализации одного из изученных методов.

В ходе работы для достижения описанной цели были поставлены следующие задачи.

1. Изучить общие понятия и терминологию в области трехмерной реконструкции.
2. Выполнить поиск и изучение существующих методов 3D-реконструкции поверхностей лиц.
3. Реализовать приложение для реконструкции трехмерных поверхностей лиц по одному входному изображению с помощью существующих методов.

# Глава 1. Обзор методов трехмерной реконструкции

Методы трехмерной реконструкции лиц различаются в зависимости от имеющихся ресурсов, поставленных требований и сфер применения. Далее в каждом параграфе представлено краткое описание общих подходов к 3D-реконструкции лиц, а также приведен обзор существующих методов, подходящих для решения поставленной задачи.

## 1.1. Основные принципы 3D-реконструкции

Все методы трехмерной реконструкции можно сгруппировать в две категории: активные и пассивные [19].

*Активные методы* взаимодействуют с восстанавливаемым объектом, механически или радиометрически. Данное взаимодействие осуществляется через излучение в направлении объектов сцены с целью извлечения дополнительной информации, на основании которой будут работать реконструирующие алгоритмы, и последующее измерение отраженной части излучения. Измерение производится посредством датчиков (сенсоров), параметры которых достаточны для восприятия излучения в объеме, необходимом для последующей обработки.

Простым примером механического подхода является применение глубиномера для измерения расстояния до объекта. Более часто применяемые радиометрические методы испускают излучение на объект, а затем замеряют его отраженную часть. Примерами могут служить движущиеся источники освещения, применение разноцветного освещения, лазерные дальномеры и LIDAR, микроволновые и ультразвуковые датчики и про-

чие методы трехмерного сканирования.

*Пассивные методы* 3D-реконструкции не оказывают влияния на восстанавливаемый объект, а лишь используют датчик для измерения излучения, отраженного или излучаемого поверхностью объекта. Обычно, в качестве сенсора применяют фоторецепторы камер, чувствительные к видимому диапазону. На вход алгоритма обработки подается либо набор из нескольких изображений (одно, два или более), либо видеопоток. В таком случае говорят о *фотограмметрии* – реконструкции, основанной на изображениях [14]. Стоит отметить, что именно такие методы реконструкции применимы для систем обеспечения безопасности, систем бесконтактного контроля, а также в рекламных и других приложениях, где необходимо реконструировать поверхность лица в режиме реального времени по фотографиям или видеоряду.

Процесс реконструкции трехмерных моделей по фотографиям в основном состоит из следующих этапов: калибровка и определение ориентации камеры, извлечение ключевых точек из изображения, генерация трехмерных облаков точек, формирование поверхности и наложение текстуры. Калибровка камеры имеет решающее значение для получения точных моделей.

Существующие на данный момент системы 3D распознавания лиц используют специальное оборудование для реконструкции трехмерной модели лица (сенсорные системы) [21]. Одной из широко применяемых в трехмерной реконструкции сенсорной системой является *стереокамеры*, которая представляет собой две (или более) идентичные камеры с параллельными оптическими осями, которые одновременно захватывают одну и ту же сцену и получают два изображения с разных углов обзора (*стереопару*) [15]. На полученных изображениях находятся ключевые точки, вычисляется положение сопоставленных точек в трехмерном пространстве и выполняется их *триангуляция* – определение трехмерных координат точек по координатам их проекций. Ключевой шаг такого процесса реконструкции – получение *карты глубин* – растрового изображения, в каждом пикселе которого содержится информация о значении глубины,

представляющего собой расстояние от точки поверхности объекта до камеры. На основании карты глубины строится трехмерная модель наблюдаемой сцены.

Основная проблема стереозрения заключается в идентификации общих точек в парах изображений, то есть решении известной проблемы соответствия. Также стерео-подход требует высокой точности калибровки камер, а при недостатке пространственной структуры объекта, отсутствии ярко выраженных характерных точек алгоритмы стереовосстановления находят только грубые детали формы объекта. Тем не менее, подход имеет такие преимущества, как простота и невысокая стоимость, отсутствие необходимости в дополнительном оборудовании (например, в специальных источниках света). Поэтому стереозрение имеет значительные применения в робототехнике и в компьютерном зрении, где суть проблемы заключается не в точном получении высококачественных данных, а в их интерпретации (например, для планирования движения, предотвращения столкновений и захвата объектов).

В качестве альтернативы стереозрению и подобным методам, использующим камеры [20], появляются методы трехмерной реконструкции, использующие для восстановления поверхности лица всего одно 2D-изображение [3], либо несколько изображений одного человека, но без наложения на них каких-либо ограничений в освещении, позе, используемой камеры и т.п. [13, 11]. Алгоритмы реконструкции лица по одному входному изображению часто основаны на использовании параметрической модели лица, введенной в [3] и подробно описанной в следующем параграфе.

Ввиду того, что стереозрение требует наличие дорогостоящего оборудования и значительных вычислительных мощностей, в ходе данной работе мы сфокусировались на реконструкции лиц с использованием одного изображения. Тем не менее мы также рассмотрели методы 3D-реконструкции лиц по стереопаре [6, 9] и видеопотоку [5, 12, 10].

## 1.2. 3D Face Morphable Model

Для достижения независимости трехмерной реконструкции от позы и освещенности лица на изображении полезно использовать 3D-информацию об объекте. По таким соображениям Blanz и Vetter почти 20 лет назад ввели понятие параметрической модели лица (3D Morphable Model (3DMM)) [3]. **Параметрическая модель лица** – это заданная параметрически трехмерная функция, основанная на линейной комбинации трехмерных моделей лиц, полученных в результате лазерного сканирования. Поверхность и текстура лица описываются соответствующими векторами коэффициентов. Таким образом, процесс реконструкции трехмерной поверхности лица с использованием параметрической модели заключается в подборе параметров, описывающих ее поверхность.

Как описано авторами модели [3], представим геометрию поверхности лица с помощью вектора формы  $S = (X_1, Y_1, Z_1, X_2, \dots, Y_n, Z_n)^T \in \mathfrak{R}^{3n}$ , который содержит  $X, Y, Z$  координаты  $n$  вершин поверхности. Текстуру лица представим в виде вектора текстуры  $T = (R_1, G_1, B_1, R_2, \dots, G_n, B_n)^T \in \mathfrak{R}^{3n}$ , который содержит  $R, G, B$  значения цвета  $n$  соответствующих вершин поверхности лица. Параметрическая модель лица построена с использованием базы данных из  $m$  экземпляров лиц, каждое из которых представлено своими вектором формы  $S_i$  и вектором текстуры  $T_i$ .

Полагая, что все лица состоят в полном соответствии точек, то есть вершины с одинаковым индексом в сетке соответствуют одной и той же семантической точке на каждом лице, новые формы  $S_{model}$  и текстуры  $T_{model}$  могут быть выражены в барицентрических координатах как линейная комбинация форм и текстур  $m$  лиц-экземпляров базы данных:

$$S_{mod} = \sum_{i=1}^m a_i S_i, \quad T_{mod} = \sum_{i=1}^m b_i T_i, \quad \sum_{i=1}^m a_i = \sum_{i=1}^m b_i = 1. \quad (1. .1)$$

Таким образом, произвольное новое лицо может быть сгенерировано с помощью варьирования параметров формы  $\vec{a}$  и текстуры  $\vec{b}$  лица.

Для оценки реалистичности сгенерированного лица при составле-



нии параметрической модели лица авторами статьи [3] была произведена оценка вероятностного распределения коэффициентов  $a_i$  и  $b_i$ . Это распределение позволяет контролировать вероятность коэффициентов и, следовательно, регулирует вероятность появления конкретного лица.

Считая, что лица в базе данных распределены нормально, с помощью метода главных компонент (РСА), можно выполнить преобразование к ортогональной системе координат:

$$S_{model} = \bar{S} + \sum_{i=0}^{m-1} \alpha_i s_i, \quad T_{model} = \bar{T} + \sum_{i=0}^{m-1} \beta_i t_i$$

где  $\bar{S}$ ,  $\bar{T}$  – средние форма и текстура лица параметрической модели,  $s_i$ ,  $t_i$  – собственные векторы ковариационных матриц форм и текстур,  $\alpha_i$ ,  $\beta_i \in \Re^{m-1}$  – коэффициенты формы  $S_{model}$  и текстуры  $T_{model}$  лица соответственно.

Выразительность модели может быть увеличена с помощью разделения лиц на сегменты, такие как глаза, нос, рот и окружающая область (см. рис. 1. .1), которые при трансформациях изменяются независимо друг от друга. Такая сегментация эквивалентна разбиению векторного пространства лиц на независимые подпространства. Таким образом, полная 3D модель лица может быть сгенерирована путем соединения всех сегментов, полученных по отдельности с помощью линейной комбинации.



Рис. 1. .1: Сегменты лица, обрабатываемые независимо: глаза, нос, рот, окружающая область. [11]

## *Сопоставление параметрической модели с изображениями лиц*

Ключевым элементом работы [3] является алгоритм для автоматического сопоставления параметрической модели с одним или несколькими изображениями. В этом алгоритме наряду с описанными параметрами формы и текстуры лица используются также параметры рендеринга. Вектор параметров рендеринга  $\vec{r}$  описывает такие показатели, как положение камеры (азимут и высоту), масштаб объекта, поворот и сдвиг плоскости изображения, интенсивность окружающего и направленного света и др.

В ходе алгоритма коэффициенты 3D-модели оптимизируются вместе с набором параметров рендеринга так, что они создают изображение как можно более близкое к входному изображению в терминах евклидового расстояния

$$E_I = \sum_{x,y} \|I_{input}(x,y) - I_{model}(x,y)\|^2. \quad (1. .2)$$

Алгоритм начинает итеративную оптимизацию от среднего лица и с грубо оцененными пользователем параметрами рендеринга. Далее в цикле создается текстура, отображенная на 3D-поверхность лица, по текущим параметрам модели, визуализируется изображение и обновляются параметры в соответствии с остаточной разницей.

Наличие многочисленных статей, посвященных реконструкции лиц с использованием параметрической модели, подтверждают тот факт, что с появлением 3DMM анализ изображений лиц вышел на новый уровень. Однако следует отметить, что основное препятствие в использовании 3DMM заключается в трудоемком процессе построения модели – создании собственной базы трехмерных лиц, что требует точного и быстрого 3D-сканера, сканировании нескольких сотен личностей и расчете плотного соответствия точек между полученными трехмерными моделями.

В связи с этим, была разработана и описана [1] общедоступная параметрическая модель *Basel Face Model* (BFM), бесплатная для использования в некоммерческих целях. Модель BFM построена на отсканированных

лицах 100 мужчин и 100 женщин в нейтральном выражении. Параметрическая модель лица BFM содержит 53490 вершин, составляющих 160470 плоскостей.

Эта модель не только позволяет разрабатывать алгоритмы анализа изображений на основе 3DMM, но также включает новые функции, как например генерация синтетического лица по случайным коэффициентам модели (рис. 1. .2). Функции работы с общедоступной моделью реализованы на языке Matlab.

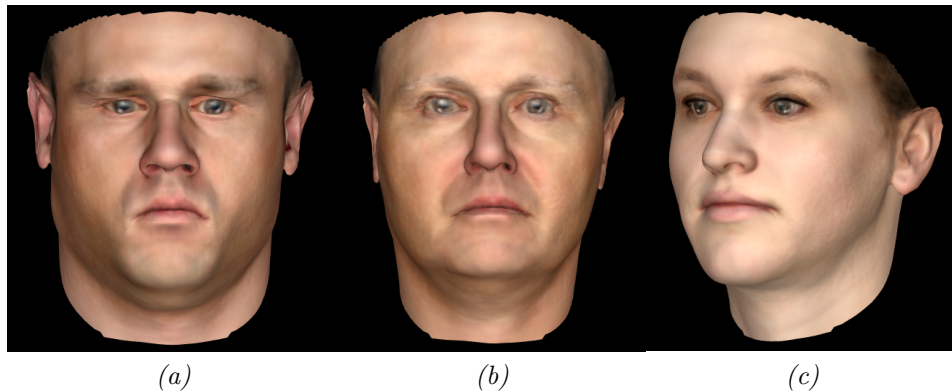


Рис. 1. .2: Синтетические лица, автоматически сгенерированные по случайным параметрам модели BFM.

### 1.3. Меры качества реконструкции для подгонки к изображениям

Процесс получения реконструкции лица наиболее близкого к исходному изображению называется *подгонкой* (fitting). Подгонка изображения выполняется с использованием какой-либо меры качества реконструкции, или иначе меры близости реконструированной поверхности с исходным изображением.

Сопоставлений изображений с помощью евклидова расстояния по формуле 1. .2 недостаточно, поскольку они нечувствительны к небольшим, но важным ошибкам и артефактам, из-за которых реконструкция лица становится ненатуральной по мнению человека.

Для людей качество может означать, насколько естественным и правдоподобным выглядит трехмерное лицо, а также насколько он похож на человека на изображении. Для неудачных реконструкций оба критерия, как правило, нарушаются одновременно, поэтому большинство из методов автоматической оценки качества реконструкции обычно оценивают только правдоподобие и не измеряют сходство с входным лицом. Целью исследований [11] было выявление меры качества, наиболее близкой к оценкам, которые присваивают реконструкциям люди.

Более того, авторы статьи утверждают, что во многих реальных приложениях доступно более одной фотографии человека, поэтому автоматизированный алгоритм реконструкции может использовать избыточные данные из нескольких изображений для увеличения надежности и робастности.

Предложенный авторами [11] алгоритм производит отдельные трехмерные реконструкции для каждого входного изображения, а затем объединяет их в оптимальное общее решение. Метод отдельно работает на разных участках лица, выбирает наиболее правдоподобные реконструкции и объединяет их в одно трехмерное лицо.

Ключевой компонент алгоритма – новый показатель визуального качества трехмерных реконструкций, основанный на поверхностных нормалях. Автоматическая оценка качества изображения в компьютерной графике и компьютерном зрении является фундаментальной задачей. Простых сопоставлений изображений недостаточно, поскольку они нечувствительны к небольшим, но важным ошибкам и артефактам.

Было установлено, что потенциальные причины неудачных реконструкций, такие как световые эффекты, окклюзии или экстремальные выражения лица, обычно приводят алгоритм к моделям, далеким от средней, и очень чувствительной мерой для этого является отклонение нормалей поверхности от среднего.

Оценка качества реконструкций с использованием поверхностных нормалей определяется следующим образом. Полагая, что параметрическая модель, определенная уравнениями 1.1, приведена к полному поточечно-

му соответствию между вершинами  $i$  модели, новая мера качества  $d_{normal}$  анализирует разность между нормальными  $n_i$  поверхности реконструированного лица и нормальными  $n'_i$  среднего лица:

$$d_{normal} = \frac{1}{n} \sum_{i=1}^n \arccos \frac{n_i \cdot n'_i}{\|n_i\| \cdot \|n'_i\|}. \quad (1.3)$$

Таким образом, нормальное расстояние определяется путем вычисления угла между нормальными среднего и восстановленного лица для пары соответствующих вершин. Для каждого сегмента лица (глаза, рот, нос и окружающая область) можно определить индивидуальное расстояние  $d_{normal}$ , которое отражает правдоподобие конкретного сегмента. Эти значения усредняются для получения глобального значения расстояния.

Однако стоит заметить, что у поверхностей человеческих лиц нормали в некоторых вершинах на носу, глазах или губах меняются больше, чем другие. Авторы статьи проанализировали оригинальные 200 трехмерных сканов параметрической модели и создали разные весовые карты, которые учитывают эти локальные различия путем масштабирования областей с высокой вариацией нормалей. Затем была найдена наилучшая карта весов, которая определяется с помощью выражения  $\omega = 1 - \frac{\bar{\phi}_i - \bar{\phi}_{min}}{\bar{\phi}_{max} - \bar{\phi}_{min}}$ , где  $\bar{\phi}$  – средний угол отклонения нормали  $n_i$  от средней нормали  $n'_i$  в каждой вершине  $i$  на всех лицах базы данных, и взвешенное нормальное расстояние определяется выражением:

$$d_{normalW} = \frac{1}{n} \sum_{i=1}^n \omega_i \arccos \frac{n_i \cdot n'_i}{\|n_i\| \cdot \|n'_i\|}. \quad (1.4)$$

Кратко общий алгоритм можно описать следующими шагами:

- выполнить подгонку 3DMM для каждого из входных изображений человека,
- выбрать лучшие результаты на каждом сегменте лица с помощью  $d_{normal}$ ,

- вычислить взвешенные линейные комбинации сегментов,
- объединять их в одно трехмерное лицо.

Реконструируемая поверхность каждого сегмента определяется взвешенной линейной комбинацией соответствующих сегментов, основанная на порядке реконструкций сегментов в списке ранжирования. Вес сегмента уменьшается с его рангом. Таким образом, комбинированная поверхность для каждого отдельного сегмента

$$S_{seg} = \sum_{i=0}^{m-1} \alpha_i S_{seg,i}$$

определяется  $m$  индивидуальными реконструкциями соответствующих сегментов  $S_{seg,i}$ , взвешенных с помощью

$$\alpha_i = \frac{1 - (i \cdot \frac{1}{m})}{\sum_{c=0}^{m-1} 1 - (c \cdot \frac{1}{m})}.$$

По результатам экспериментов, авторы алгоритма также определили пороговое значение меры качества, который отделяет правдоподобные от неправдоподобных реконструкций и определяет, какие реконструкции рассматриваются во взвешенной сумме, а какими стоит пренебречь.

Автоматизированная 3D-реконструкция, предложенная в этой статье, компенсирует низкую точность и надежность автоматически определяемых позиций, используя более одного изображения лица. Стоит заметить, что, в отличие от стерео и многовидовых алгоритмов, в данном подходе учитываются нежесткие деформации мимики и большие различия в условиях формирования изображения.

Экспериментальные сравнения существующих мер качества с оценками, которые устанавливали люди тестируемым реконструкциям, показали, что новые мера качества, основанные на поверхностных нормалях (1. .3), (1. .4) значительно превосходят ранее используемые оценки качества реконструкции.

## Глава 2. Реализация

В данном разделе описана реализация программы, производящей трехмерную реконструкцию лиц по одному входному изображению.

### 2.1. Используемые технологии

В рамках данной работы для реализации трехмерной реконструкции лиц по фотографиям была использована библиотека *eos* с открытым исходным кодом [17], созданная для подгонки 3D-моделей с помощью параметрической модели лица (1. .1), написанная на современном C++11/14 и использующая в своей реализации популярную открытую библиотеку компьютерного зрения OpenCV. Язык C++ традиционно используется при решении задач компьютерного зрения в связи с тем, что программы на C++ отличается быстроедействие и производительностью, совместимы с различными устройствами и легко встраиваются во всевозможные системы, а с появлением новых стандартов язык стал более простым в изучении и чтении.

### 2.2. Библиотека eos

Библиотека eos предоставляет базовые функциональные возможности для использования моделей лиц и подгонки к параметрической модели.

На данный момент библиотека имеет следующий функционал:

- классы *MorphableModel* и *PcaModel* для представления 3DMM с базовыми функциями;

- параметрическая модель Surrey Morphable Face Model содержащая только форму лица с низким разрешением, созданная на основе трехмерных сканирований лиц;
- быстрая линейная подгонка позы, формы и выражения лица;
- извлечение текстуры Isomap для получения независимого от позы представления о текстуре лица ;
- Python и Matlab интерфейсы для работы с библиотекой;
- подгонка формы, камеры и цветовой модели с помощью нелинейной функцией стоимости с использованием библиотеки Ceres (экспериментальный функционал).

Также доступен исходный код [2] программы для просмотра параметрической модели лица, варьирования параметров формы и выражения лица.

Параметрическая модель Surrey 3D Morphable Face Model [18] построена на 170 трехмерных моделях, полученных в результате лазерного сканирования лиц людей различной этнической принадлежности. Последнее позволяет модели хорошо представлять лица людей различных национальностей, в то время как модель BFM ориентирована в основном на лица европейской внешности.

Параметрическая модель Surrey имеет три уровня разрешения: 3448, 16759 и 29587 вершин. В библиотеке по умолчанию используется модель с наименьшим разрешением для легкости и быстроты использования, другие модели доступны по лицензии университета Surrey.

## 2.3. Описание алгоритма

Алгоритм трехмерной реконструкции библиотеки eos в качестве входных данных использует одно изображение лица и список ключевых точек (ibug landmarks<sup>1</sup>). В ходе данной работы был использован оригинальный

---

<sup>1</sup><https://ibug.doc.ic.ac.uk/resources/facial-point-annotations/>



алгоритм детекции ключевых точек с помощью библиотеки dlib.

Результат реконструкции сохраняется в файл формата .obj, который является общепринятым среди разработчиков приложений 3D-графики. Использование данного формата позволяет открывать модели, полученные в результате реконструкции, в большинстве программ для работы с трехмерными объектами, что особенно важно для сферы компьютерных игр, 3D-фильмов и виртуальной реальности.

- Первый шаг алгоритма представляет собой *определение положения камеры*. Учитывая расположение ключевых точек и известное их соответствие в параметрической модели, производится оценка позы лица (или положение камеры, что в данном случае является эквивалентной задачей). Для этого обнаруженные ключевые точки на изображении и соответствующие трехмерные точки модели нормализуются с помощью преобразований подобия, которые переносит центроид изображения и модели и производит масштабирование. После этого вычисляется матрица камеры, что предполагает решение линейной системы уравнений. Этот процесс очень быстрый, и для вычисления требуется всего несколько миллисекунд.
- Следующий шаг алгоритма заключается в *реконструкции трехмерной формы* лица с использованием полученной матрицы камеры. Восстановление трехмерной поверхности реализовано с помощью подгонки формы к ключевым точкам, путем нахождения наиболее вероятного вектора коэффициентов формы параметрической модели, минимизируя следующую функцию стоимости:  $E = \sum_{i=1}^{3N} \frac{(y_{m2D,i} - y_i)^2}{2\sigma_{2D}^2} + \|\alpha\|_2^2$ , где  $N$  – количество ключевых точек,  $y$  – обнаруженные двумерные ключевые точки в однородных координатах,  $\sigma$  – отклонение для ключевых точек, а  $y_{m2D,i}$  – проекция ключевых точек формы 3D Morphable Model на двумерную плоскость с использованием вычисленной матрицы камеры. Оценка формы выполняется так же быстро, как и оценка позы и выполняется несколько миллисекунд.
- После получения позы и формы лица существует плотное соответствие

между вершинами полученной трехмерной сетки и лицом на входном изображении. Тогда выполняется *извлечение текстуры* из входного изображения, ее сохранение, и затем текстура может быть повторно отображена на модели в произвольных позах. На рис. 2. .1 показан пример подгонки формы и модели камеры с извлеченной текстурой лица. На текстурной карте области окклюзии изображены как белые пятна, однако в isomap они идентифицируются альфа-каналом.

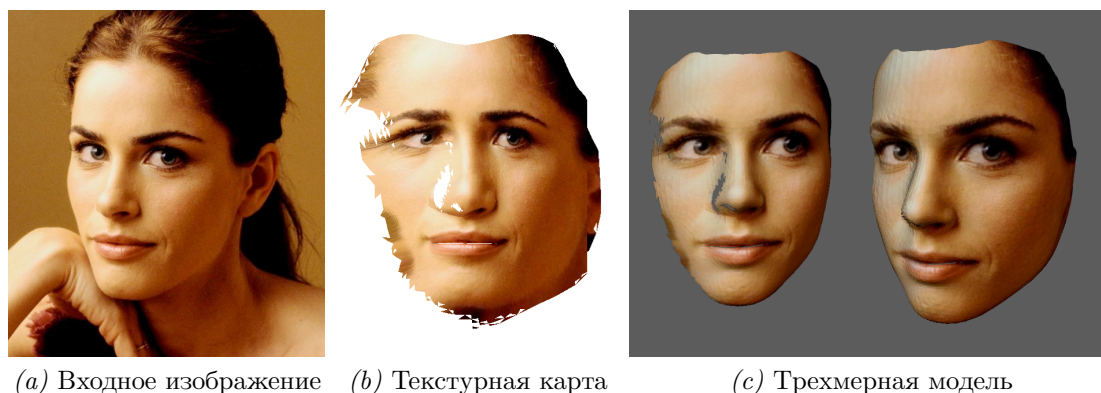


Рис. 2. .1: Пример результата работы алгоритма

В этом разделе было описаны возможности библиотеки eos для решения поставленной задачи. По оценкам авторов библиотеки [18] определение позы и ориентировочная подгонка трехмерной модели выполняются за миллисекунды, в то время как переотображение текстуры является самым медленным компонентом алгоритма, занимающий около 100 миллисекунд (на процессоре Intel Core i7-4700MQ). Отказ от использования дополнительных методов ускорения алгоритма обуславливается уменьшением зависимостей фреймворка и его независимостью от платформ.

## Глава 3. Эксперименты

Для тестирования реализации и проведения экспериментов были использованы фотографии из базы данных MUCT [8]. Одним из преимуществ данной базы изображений является наличие фотографий людей различных национальностей, что позволяет проверить ориентированность параметрической модели Surrey на изображения людей различной этнической принадлежности.

Первое, на что стоит обратить внимание при анализе результатов работы алгоритма – это ”пробелы” в текстурных картах (см. рис. 2. .1 (b)). Как уже отмечалось, это области окклюзии (сокрытия), то есть такие области, которые скрыты от камеры в силу того, что выступающие части лица частично закрывают видимость других его частей. Так как в данном алгоритме текстура извлекается непосредственно из входного изображения и не производится ее подгонка с помощью параметрической модели, наличие таких ”пробелов” в текстурной карте очевидно. Тем не менее, реконструированная поверхность лица является цельной и не имеет артефактов.

На фотографиях использованной базы в основном присутствуют лица в положении близком к ”анфас”, поэтому текстурная карта имеет лишь незначительные ”пробелы” (см. рис. 3. .1).

Далее стоит отметить, что в связи с тем, что в библиотеке eos используется параметрическая модель низкого разрешения, видно, что реконструированная поверхность не является гладкой, а состоит из маленьких полигонов. В приложениях, где необходима высокая точность 3D-реконструкции лиц, такого низкополигонального представления модели может быть недостаточно. Поэтому для получения более гладкой поверх-

ности можно использовать параметрическую модель с большим разрешением, либо увеличивать количество полигонов полученной модели методами сглаживания и др. Однако во втором случае точность реконструкции может несколько снизиться.

Оценить качество реконструированной поверхности можно программными средствами путем сравнения с результатами лазерного сканирования или других алгоритмов автоматической реконструкции. В связи с отсутствием трехмерных моделей используемой базы лиц, попытаемся провести оценку качества полученных моделей визуально.

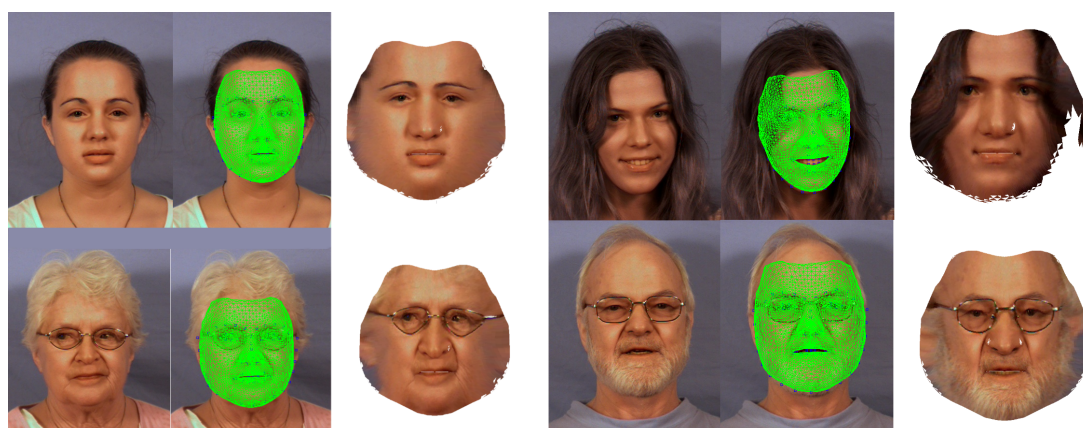


Рис. 3. .1: Результаты работы алгоритма

Наложение трехмерной сетки на входное изображение (см. рис. 3. .1) демонстрирует подгонку модели к изображению. Однако при детальном рассмотрении полученных моделей можно заметить существенные артефакты. На рисунке 3. .2 хорошо видны следующие недостатки:

- отсутствуют ярко выраженные носогубные морщины;
- для лиц с широким носом текстура носа выходит за его пределы на трехмерной модели;
- 3D модель не учитывает асимметрию лица (кривизну носа, неровность линии глаз).

Данные артефакты свидетельствуют, что проблема алгоритма заключается в подгонке 3D-модели к входному изображению. Для улучше-



Рис. 3. .2: Артефакты подгонки трехмерной модели к изображениям

ния качества реконструкции использованного алгоритма можно предложить следующие решения: уточнение местоположения ключевых точек на лице, использование нелинейной меры качества, например (1. .3), при подгонке модели к изображению.

Таким образом, экспериментальные исследования показали, что алгоритм трехмерной реконструкции, реализованный в библиотеке eos имеет ряд достоинств (быстрота, простота использования) и ряд недостатков (отсутствие текстуры в параметрической модели, появление артефактов при подгонке моделей). Исходя из этого можно сделать вывод о том, что данный алгоритм применим для задач трехмерной реконструкции, не требующих высокой точности, и не располагающих большими ресурсами, например, в сфере компьютерных игр, виртуальной реальности и рекламы. А в таких сложных задачах как трехмерное распознавание лиц для систем безопасности или контроля следует выбрать более точный и, возможно, более ресурсоемкий метод реконструкции лиц.

## Заключение

В ходе данной работы были изучены основные принципы трехмерной реконструкции лиц. Наибольшее внимание было уделено методам 3D-реконструкции с использованием параметрической модели лица, которые работают всего по одному входному изображению, что позволяет их использовать в условиях ограниченных ресурсов и без дополнительного оборудования.

Для решения поставленной задачи было реализовано приложение, выполняющее трехмерную реконструкцию лиц по одному входному изображению с использованием библиотеки eos. В ходе экспериментальных исследований были выявлены недостатки данного подхода, такие как появление артефактов на текстуре и несовпадения черт реконструированного лица с изображенным на фотографии. Появление данных артефактов связано с использованием параметрической модели Surrey и линейной функции оценки качества реконструкции при подгонке изображений. Для улучшения результатов реконструкции с помощью библиотеки eos можно предложить использовать нелинейную меру качества при подгонке моделей, а также уточнить поиск ключевых точек на изображении.

## Литература

- [1] A 3D Face Model for Pose and Illumination Invariant Face Recognition / IEEE. — Genova, Italy, 2009.
- [2] 3D model viewer for the eos Morphable Model library. — <https://github.com/patrikhuber/eos-model-viewer>. — Accessed: 2017-05-20.
- [3] Blanz Volker, Vetter Thomas. A morphable model for the synthesis of 3D faces / ACM Press/Addison-Wesley Publishing Co. — 1999. — P. 187–194.
- [4] Detailed 3D reconstruction of large-scale heritage sites with integrated techniques / Sabry F El-Hakim, J-A Beraldin, Michel Picard, Guy Godin // IEEE Computer Graphics and Applications. — 2004. — Vol. 24, no. 3. — P. 21–29.
- [5] Jeni László A, Cohn Jeffrey F, Kanade Takeo. Dense 3D face alignment from 2D videos in real-time / IEEE. — Vol. 1. — 2015. — P. 1–8.
- [6] Kemelmacher-Shlizerman Ira, Basri Ronen. 3D face reconstruction from a single image using a single reference face shape // IEEE transactions on pattern analysis and machine intelligence. — 2011. — Vol. 33, no. 2. — P. 394–405.
- [7] Lanitisa Andreas, Stylianoub Georgios. RECONSTRUCTING 3D FACES IN CULTURAL HERITAGE APPLICATIONS. — 2008.
- [8] Milborrow Stephen, Morkel John, Nicolls Fred. The MUCT landmarked face database // Pattern Recognition Association of South Africa. — 2010. — Vol. 201, no. 0.

- [9] Noniterative 3D face reconstruction based on photometric stereo / Sang-Woong Lee, Patrick SP Wang, Svetlana N Yanushkevich, Seong-Whan Lee // International Journal of Pattern Recognition and Artificial Intelligence. — 2008. — Vol. 22, no. 03. — P. 389–410.
- [10] Park Unsang, Jain Anil K. 3D face reconstruction from stereo video / IEEE. — 2006. — P. 41–41.
- [11] Piotraschke Marcel, Blanz Volker. Automated 3d face reconstruction from multiple images using quality measures. — 2016. — P. 3418–3427.
- [12] Real-time high-fidelity facial performance capture / Chen Cao, Derek Bradley, Kun Zhou, Thabo Beeler // ACM Transactions on Graphics (TOG). — 2015. — Vol. 34, no. 4. — P. 46.
- [13] Roth Joseph, Tong Yiyang, Liu Xiaoming. Unconstrained 3D face reconstruction. — 2015. — P. 2606–2615.
- [14] Sansoni Giovanna, Trebeschi Marco, Docchio Franco. State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation // Sensors. — 2009. — Vol. 9, no. 1. — P. 568–601.
- [15] Sansoni Giovanna, Trebeschi Marco, Docchio Franco. State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation // Sensors. — 2009. — Vol. 9, no. 1. — P. 568–601.
- [16] VOCORD FaceControl 3D. — <http://old.vocord.ru/catalog/products/sistemy-videonablyudeniya/vocord-facecontrol-3d/>. — Accessed: 2017-05-20.
- [17] eos: A lightweight header-only 3D Morphable Face Model fitting library in modern C++11/14. — <https://github.com/patrikhuber/eos>. — Accessed: 2017-05-20.



- [18] A multiresolution 3D Morphable Face Model and fitting framework / Patrik Huber, Guosheng Hu, Rafael Tena et al. // International Conference on Computer Vision Theory and Applications (VISAPP). — 2016. — P. 1–8.
- [19] Алдошкин Д. Н. ТРЁХМЕРНАЯ РЕКОНСТРУКЦИЯ. ОБЛАСТИ ПРИМЕНЕНИЯ. МЕТОДЫ И ПОДХОДЫ. — 2012.
- [20] Биометрические технологии компании "Вокорд" / Дмитрий Николаевич Заварикин, Алексей Андреевич Кадейшвили, Светлана Викторовна Коробкова et al. // Вестник Московского государственного технического университета им. НЭ Баумана. Серия «Приборостроение». — 2011. — no. 5.
- [21] Дьяченко А. В. Задача 3D распознавания лиц: современные методы решения // Штучний інтелект. — 2011.
- [22] Ершаков К. Продвинутые технологии распознавания. Развитие 3D-идентификации и сканирования лица // Системы безопасности. — 2014. — no. 1.
- [23] Заварикин Д. Н. Принципы построения и преимущества системы 3D-распознавания лиц // Технология защиты. — 2010. — no. 4.
- [24] Щеголева Н. Восстановление 3D поверхности лица // GraphiCon'2013. — 2013. — P. 135–139.